Contents

1 Interval estimation

Confidence

It is not always satisfiable with a single guess on the value of an unknown parameter θ.

In stead we use *confidence sets*, fx. confidence intervals, or confidence ellipsoides (in two dimensions).

Suppose that we have a sample $X_1, \dots, X_n$ with a distribution depending on an unknown and real value parameter θ.

A confidence interval consists of two limits:

- a lower limit

  $C_1 = g_1(X_1, \dots, X_n)$ and

- an upper limit $C_2 = g_2(X_1, \dots, X_n)$.

The interval $[C_1, C_2]$ has *degree of confidence* γ, if

$$P(C_1 \le \theta \le C_2) = \gamma.$$

We want as *small* intervals with as *high* degree of confidence as possible!

1.1 Confidence interval for the normal mean - variance known

Confidence interval for the normal distribution - variance known

If the variance is *known* equal to the priori variance σ_0^2, then

- $Z = \frac{\sqrt{n}(\bar{X}-\mu)}{\sigma_0}$ has a standard normal distribution

which e.g. means that $P(-1.96 < Z < 1.96) = 0.95$, which can be transformed to an interval for μ.

In general, a confidence interval with degree $\gamma = 1 - \alpha$ for μ is given by

$$c_1 = \bar{x} - z_{1-\alpha/2}\sigma_0/\sqrt{n}, \quad c_2 = \bar{x} + z_{1-\alpha/2}\sigma_0/\sqrt{n},$$

where $z_{1-\alpha/2}$ is the $(1-\alpha/2)$-quantile in the standard normal distribution, i.e. $\Phi(z_{1-\alpha/2}) = 1 - \alpha/2$.

1.2 Confidence interval for the normal mean - variance unknown

Confidence interval for the normal distribution - variance unknown

If the variance is unknown and estimated by the posterior variance s^2 with $d = n - 1$ redundants, we still focus on the standardized variable, but where σ_0 is substituted by s

- $T = \frac{\sqrt{n}(\bar{X}-\mu)}{s}$

This is not normal as we have introduced some extra variability by using s instead of σ_0.

The distribution of T is called the t-distribution with d degrees of freedom. Then the confidence interval with degrees $\gamma = 1 - \alpha$ for μ is

$$c_1 = \bar{x} - t_{1-\alpha/2}(d)s/\sqrt{n}, \quad c_2 = \bar{x} + t_{1-\alpha/2}(d)s/\sqrt{n}.$$

where $t_{1-\alpha/2}$ is the $(1-\alpha/2)$-quantile in the standard normal distribution, i.e. $P(T < t_{1-\alpha/2}) = 1 - \alpha/2$.

1.3 Student's t-distribution

Student's t

or the t-*distribution* with d degrees of freedom is formally derived as

$$T = \frac{Z}{\sqrt{Y/d}},$$

where Z and Y are independent, Z is standard normal, and Y χ^2-distributed with d degrees of freedom.

The density function is

$$f_T(t; d) = c_d \left(1 + t^2/d\right)^{-(d+1)/2}.$$

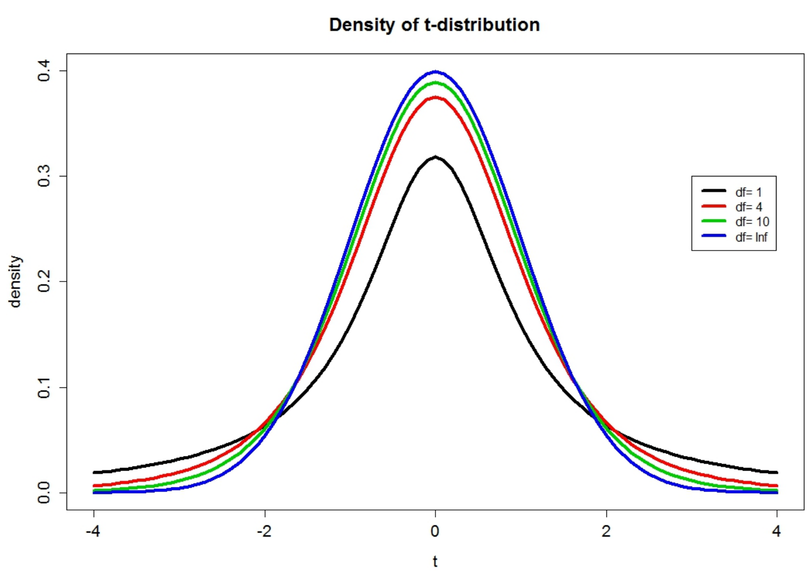
Student's t

Heavy tails for small d.

For $d \to \infty$ the t-distribution approaches a standard normal distribution. Difficult to distinguish for $d > 30$.

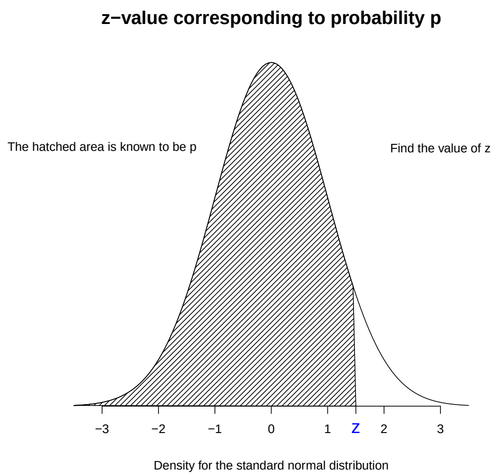
"Student" is a pseudonym for W. S. Gossett, who invented this distribution around year 1900.

Examples of t-distributions



1.4 Quantiles of the standard normal and t-distributions

Quantiles of the standard normal and t-distributions



May be determined using python

```
>>> from scipy.stats import norm
>>> norm.ppf([.95,.975,.99,.99865])
array([ 1.64485363,  1.95996398,  2.32634787,  2.99997699])
```

Similarly for the t-dist. with df=2

```
>>> from scipy.stats import t
>>> t.ppf([.95,.975,.99,.99865],2)
array([  2.91998558,   4.30265273,   6.96455672,  19.20601589])
```

Example again

Observations: $119, 112, 114$, i.e. $\bar{x} = 115$, $s = 3.61$.

95% confidence interval when known standard deviation equals 4:

$$\bar{x} \pm 1.96 \times 4/\sqrt{3} = 115 \pm 1.96 \times 4/\sqrt{3} = 115 \pm 4.53,$$

since $z_{0.975} = 1.96$.

95% confidence interval with unknown standard deviation:

$$\bar{x} \pm 4.30 \times s/\sqrt{3} = 115 \pm 4.30 \times 3.61/\sqrt{3} = 115 \pm 8.97,$$

since $t(2)_{0.975} = 4.30$. Wider due to the extra uncertainty on the standard deviation.

1.5 Confidence interval for the normal standard deviation

Confidence interval for the standard deviation

Let the posterior variance s^2 be estimated with d redundants, then a confidence interval with degree $\gamma = 1 - \alpha$ for σ has limits

$$c_1 = \sqrt{\frac{d}{\chi^2(d)_{1-\alpha/2}}}s, \quad c_2 = \sqrt{\frac{d}{\chi^2(d)_{\alpha/2}}}s.$$

Here $\chi^2(d)_\lambda$ is the λ-quantile in the χ^2-distribution with d degrees of freedom.

Example again

Observations: $119, 112, 114$, i.e. $\bar{x} = 115$, $s = 3.61$.

95% confidence interval for the standard deviation:

$$\left[\sqrt{\frac{2}{\chi^2(2)_{.975}}}3.61, \sqrt{\frac{2}{\chi^2(2)_{.025}}}3.61\right],$$

Python:

```
>>> from scipy.stats import chi2
>>> chi2.ppf([.025,.975],2)
array([ 0.05063562,  7.37775891])
```

Since $\chi^2(2)_{.975} = 7.38$ and $\chi^2(2)_{.025} = 0.051$ we acheive the confidence interval $[1.88, 22.6]$

Very wide - be careful when using posterior variances with few redundants!!

1.6 Summary of confidence intervals

Summary of $(1-\alpha)$-confidence intervals

Given a sample $x_1, \ldots, x_n$ from $\mathcal{N}(\mu, \sigma^2)$, let $\hat{\mu} = \bar{x} = \frac{x_1 + \cdots + x_n}{n}$ and $\hat{\sigma}^2 = s^2 = \frac{1}{n-1}\sum(x_i - \bar{x})^2$.

- σ has a known value σ_0. Confidence interval for μ:

  $$c_1 = \bar{x} - z_{1-\alpha/2}\sigma_0/\sqrt{n}, \quad c_2 = \bar{x} + z_{1-\alpha/2}\sigma_0/\sqrt{n},$$

  where $z_{1-\alpha/2}$ is the $(1-\alpha/2)$-quantile in the standard normal distribution.

- The value of σ is unknown. Confidence interval for μ:

$$c_1 = \bar{x} - t_{1-\alpha/2}(d)s/\sqrt{n}, \quad c_2 = \bar{x} + t_{1-\alpha/2}(d)s/\sqrt{n}.$$

  where $t_{1-\alpha/2}(d)$ is the $(1-\alpha/2)$-quantile in the t-distribution with $d = n-1$ degrees of freedom.

- Confidence interval for σ: $c_1 = \sqrt{\frac{d}{\chi^2(d)_{1-\alpha/2}}}s$ and $c_2 = \sqrt{\frac{d}{\chi^2(d)_{\alpha/2}}}s$. Here $\chi^2(d)_\lambda$ is the λ-quantile in the χ^2-distribution with $d = n-1$ degrees of freedom.

2 Theory of testing

2.1 Hypotheses

Test problems

- In a statistical test a **hypothesis** is confronted with reality by means of the observations $x_1, \ldots, x_n$.
- The hypothesis is traditionally denoted by H_0, **null hypothesis**.
- The hypothesis can be scientific or controlling.
- The hypothesis can only be **falsified** (rejected) by a statistiscal test, never accepted.
- The hypothesis is assumed to be true, then it is checked by reality(the data), to decide whether we can rely on it.

Scientific hypotheses

- Celestrial bodies are moving with earth as a center. (Galileo Galilei)
- The speed of light is infinite (Rømer)
- The earth is shaped as a ball (Laplace)

Most scientific hypothesis are formulated in order to do falsification.

Laplace was obsessed by the idea that earth is shaped as a pear. But he never successed in **falsifying** the "ball hypothesis".

2.2 Controlling hypothesis

Controlling hypotheses

- A series of measurements are claimed to have a given precision (global test). I.e no misspecifications of weights.
- A specific measurement is not due to a gross error.
- An object has not been moved.
- An object has not been deformed.
- An object satisfies specific standards.

Controlling hypotheses are most common in measurement theory.

Example:
Difference in height observed: $119, 112, 114$.

Possible hypotheses and questions:

- The manufactorer of our measuring device claims that H_0: the standard deviation of our measurements is 2 mm.
- H_0: The difference in height should be 111 mm as specified by some standard.
- Measurements from a year ago: $110, 112, 109$. Has something been moved? H_0: Nothing has been moved.
- Is 119 an outlier? H_0: The first observation is not an outlier.

2.3 Statistical testing of a hypothesis

Construction of a statistical test

1. Choose a **test statistics** $W = g(X_1, \ldots, X_n)$ that reveals deviation from the hypothesis.
2. Decide whether large or small values of W (or both) are **critical** for the null hypothesis H_0. Often an **alternative** hypothesis H_A is specified.

3. Choose a **significance level** α. Fx. $\alpha = 5\%$.

4. Determine a **critical region** K_α, such that $P(W \in K_\alpha | H_0) = \alpha$**.** The complementary set $A_\alpha = K_\alpha^c$ is called the **accept region** of the test.

5. H_0 rejected, if $w_{\text{obs}} \in K_\alpha$.

   H_0 not rejected, if $w_{\text{obs}} \in K_\alpha^c$.

2.4 One sided test

Example continued: test on standard deviation
Observations: $119, 112, 114$.

$$\begin{array}{rll} \text{Nul hypothesis} & H_0: & \sigma = \sigma_0 = 2 \\ \text{Alternative hypothesis} & H_A: & \sigma > \sigma_0 = 2 \quad \textbf{one-sided test} \end{array}$$

- Test statistics $Y = dS^2/\sigma_0^2$, under H_0: $Y \sim \chi^2(d)$. $d = n-1 = 2, s = 3.606$ and $y_{\text{obs}} = 2 \times 3.606^2/4 = 6.5$.

- Consider large values as critical, i.e. the critical region with $\alpha = 5\%$ is:

$$K_\alpha = [\chi^2(d)_{1-\alpha}, \infty[= [\chi^2(2)_{0.95}, \infty[= [5.99, \infty[.$$

- Since the test statistics $y_{\text{obs}} = 6.5$ is in the critical region, the hypothesis is rejected for $\alpha \geq 5\%$.

2.5 Two sided test

Two-sided test
Both small and large values of W are critical for H_0. The accept region is then

$$A_\alpha = [w_{\alpha/2}, w_{1-\alpha/2}],$$

therefore the two halves of the critical region

$$K_\alpha =]-\infty, w_{\alpha/2}] \cup [w_{1-\alpha/2}, \infty],$$

each has probability $\alpha/2$.

In the example it can be natural to consider a two-sided test. Then

$$A_\alpha = [\texttt{chi2inv}(0.025, 2), \texttt{chi2inv}(0.975, 2)] = [0.0506, 7.378]$$

and H_0 is therefore accepted with $\alpha = 5\%$.

2.6 Test on standard deviation in a normal sample

Normal sample: $H_0 : \sigma = \sigma_0$.
$X_1, \ldots, X_n$ sample from $\mathcal{N}(\mu, \sigma^2)$ with μ and σ unknown.
Test: $H_0 : \sigma^2 = \sigma_0^2$.
Test statistics:

$$Y = \frac{(n-1)S^2}{\sigma_0^2}.$$

Two-sided test, i.e. alternative hypothesis $H_A : \sigma^2 \neq \sigma_0^2$, then the accept region is

$$A_\alpha = [\chi^2(n-1)_{\alpha/2}, \chi^2(n-1)_{1-\alpha/2}]$$

One-sided test, i.e. alternative hypothesis $H_A : \sigma^2 > \sigma_0^2$, then the accept region is

$$A_\alpha = [-\infty, \chi^2(n-1)_{1-\alpha}]$$

2.7 Global test

Global test.
Some time during this course you will learn about least squares adjustment.

- Our observations(typically measurements of lengths and angles) are stored in the vector b. The measurement b_i has variance $\sigma_0^2 u_i$ and the measurements are independent. σ_0^2 is the unit variance and most often set to 1.
- Our unknowns (also called the elements) are stored in the vector x. It will typically be coordinates of points.
- Observation equation(linearized): $b - b_0 = A(x - x_0) - r$ with d redundants, i.e. $d = n - p$ where n is the number of observations(length of b) and p is the number of unknown elements(length of x).

Global test.

- Estimated residual vector: $\hat{r}$
- Weight matrix: C is diagonal with $c_{ii} = \frac{1}{u_i}$

- Posterior unit variance: $s_0^2 = \frac{1}{d}\hat{r}^T C \hat{r}$

Test: $H_0 : E(s_0^2) = \sigma_0^2.$
Test statistics:

$$Y = \frac{dS_0^2}{\sigma_0^2}.$$

which in case of H_0 has a $\chi^2(d)$-distribution.

2.8 Test on mean in a normally distributed sample – known variance

Normal sample: $H_0 : \mu = \mu_0$ **when** $\sigma = \sigma_0$.
$X_1, \ldots, X_n$ sample from $\mathcal{N}(\mu, \sigma_0^2)$, μ unknown and σ_0 known. Test:

$$H_0 : \mu = \mu_0.$$

Test statistics:

$$Z = \frac{(\bar{X} - \mu_0)}{\sigma_0/\sqrt{n}}.$$

Two-sided test, i.e. alternative hypothesis $H_A : \mu \neq \mu_0$, then the accept region is

$$[z_{\alpha/2}, z_{1-\alpha/2}].$$

Example continued: test on mean – known variance

Nul hypothesis $H_0 : \quad \mu = 111$
Alternative hypothesis $H_A : \quad \mu \neq 111$ **two-sided test**

- Test statistics:

$$Z = \frac{\bar{X} - \mu_0}{\sigma_0/\sqrt{n}}; \quad z_{\text{obs}} = \frac{115 - 111}{2/\sqrt{3}} = 3.464.$$

- Both small and large values are critical, i.e. the accept region is:

$$A_\alpha = [z_{\alpha/2}, z_{1-\alpha/2}]$$

The hypothesis is rejected even for $\alpha = 1\%$, where the accept region is $A_{0.01} = [u_{0.005}, u_{0.995}] = [-2.58, 2.58]$.

2.9 Test on mean in a normally distributed sample – unknown variance

Normal sample: $H_0 : \mu = \mu_0$ **when** σ **unknown.**
$X_1, \ldots, X_n$ sample from $\mathcal{N}(\mu, \sigma^2)$, μ and σ unknown.
Test:

$$H_0 : \mu = \mu_0.$$

Test statistics:

$$T = \frac{(\bar{X} - \mu_0)}{s/\sqrt{n}}.$$

Two-sided test, i.e. alternative hypothesis $H_A : \mu \neq \mu_0$, then the accept region is

$$[t(n-1)_{\alpha/2}, t(n-1)_{1-\alpha/2}].$$

Example continued: test on mean – unknown variance
If we do not believe in the prior standard deviation 2 we use the posterior standard deviation s and the t-distribution:

- Test statistics:

$$T = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}; \quad t_{\text{obs}} = \frac{115 - 111}{3.606/\sqrt{3}} = 1.922.$$

- Both small and large values are critical, i.e. the accept region is:

$$A_\alpha = [t(d)_{\alpha/2}, t(d)_{1-\alpha/2}].$$

The hypothesis is not rejected even for $\alpha = 10\%$, where the accept region is $A_{0.1} = [t(2)_{0.05}, t(2)_{0.95}] = [-2.92, 2.92]$.